



Applying PHIN Vocabulary Access and Distribution System and Other Standard Vocabulary Resources to the Development of a Repository of Reusable Web-based Demographic Survey Questions and Coded Responses



Sonali Bora, MD, MS^{1,3}; Victoria Wangia, MS^{2,3}; Agha N. Khan, MD, MPH¹; Tonya D. Martin, MPH¹

¹ Prevention Informatics Office, Office of the Director, NCHSTP, CDC, Atlanta, GA; ² Program Evaluation Research Branch, Divisions of HIV/AIDS Prevention, NCHSTP, CDC, Atlanta, GA; ³ Public Health Informatics Fellowship Program, OWCD, CDC, Atlanta, GA

Objectives

- To map demographic question data elements and values to existing vocabulary standards using Public Health Information Network Vocabulary Services (PHIN VS).
- To develop reusable survey questions applying the data elements and values that had been mapped to standard vocabulary.

Introduction

- Vocabulary standards enable interoperable exchange of information, facilitating flow of information across applications and agencies.
- An overwhelming number of vocabulary standards exist, and determining what standard to apply to one's specific context can be challenging. Therefore, CDC realized the need to develop the PHIN VS.
- The PHIN VS is for registration, identification, mapping, authoring, and editing of standards-based vocabularies for applications.¹
- Public Health Information Network Vocabulary Access and Distribution System (PHIN VADS) is a browser-based vocabulary management system and the PHIN recommended vocabulary domain file is a searchable document, both are part of the PHIN VS.
- mrInterview is a web-based survey development tool, developed by SPSS. The tool has functionality to store libraries of questions, enabling the questions to be reusable by survey developers using the tool.
- Developing reusable questions with standard vocabulary is desirable. This allows for interoperability, especially when data is exported to be used between applications. It also provides a means for consistent interpretation of data, hence minimizing ambiguities as data is used and shared.²

Methods

- Met with PHIN vocabulary experts to discuss project objectives and strategy.
- Identified demographic-related questions, data elements and values from surveys and databases gathered from NCHSTP, and verified findings with subject matter experts.
- If the vocabulary standard existed - Mapped data elements and values to vocabulary standards drawn from PHIN VADS (figure 1), through the PHIN recommended vocabulary domain searchable file (figure 2), and from other sources recommended by PHIN vocabulary experts.
- Submitted program-specific data elements and values to PHIN vocabulary experts for the vocabulary reconciliation and registration process.
- Formulated survey questions, applying the identified standard data elements and values.
- Input survey questions into mrInterview and created a library of demographic questions.
- Applied some of the questions from the library to a survey developed in mrInterview.

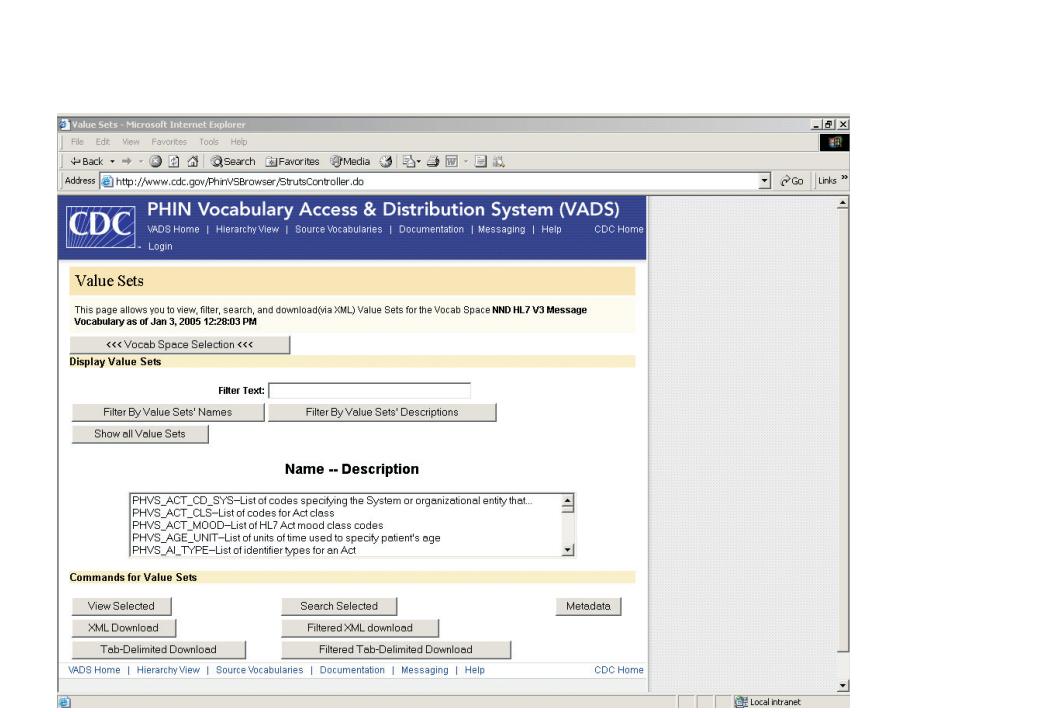


Figure 1. PHIN VADS Web Interface.

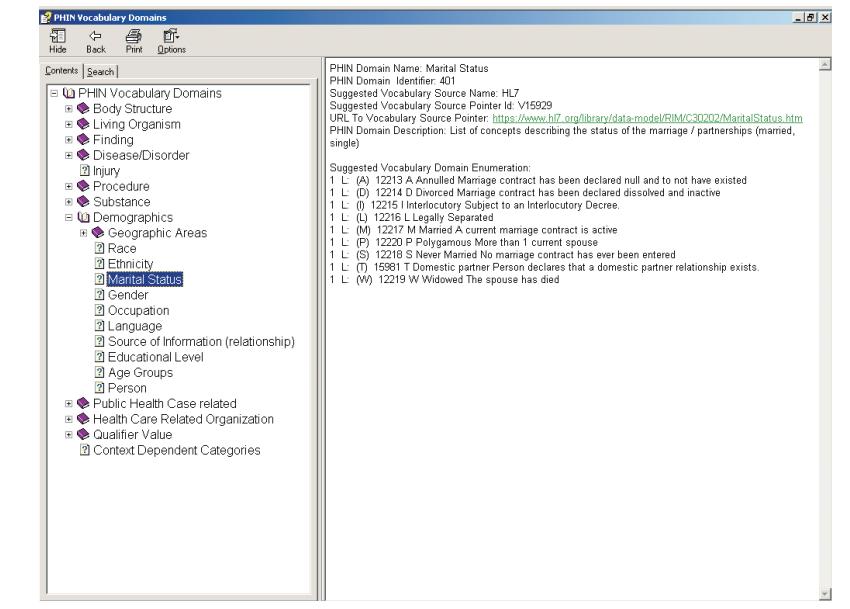


Figure 2. PHIN Recommended Vocabulary.

Results

Table 1. Vocabulary Standards Applied in The Project and Examples of Mapped Demographic Data Elements or Values.

Standard	Examples of mapped data elements/values
PHIN VADS:Bioterrorism / Laboratory Response Network (BT/LRN): ISO 3166	Country
PHIN VADS: NEDSS base system v. 1.1.3 standard reference tables	Age Unit; Birth Place
PHIN VADS: NEDSS base system 1.0 messaging codes	Sex
PHIN VADS: NND HL7 V3 message vocabulary	Insurance Type; Responses (not exhaustive: Yes, No, Unknown)
PHIN Recommended Vocabulary Searchable File:HL7 Vocabulary Domain	Ethnicity; Ethnicity Category; Race; Race Category (CDC codes submitted to HL7); Marital Status and Education
PHIN Recommended Vocabulary Searchable File:US Department of Labor (Bureau of Labor Statistics-SOC Groups) and North American Industry Classification System (NAICS) Codes	Occupations
PHIN Recommended Vocabulary Searchable File:ISO - 639	Language
PHIN Recommended Vocabulary Searchable File:SNOMED CT	Person eg. individual, infant, beggar, caregiver; Relative; Gender
PHIN Recommended Vocabulary Searchable File:FIPS Publications	City, County, State, Metropolitan Areas
PHIN Recommended Vocabulary Searchable File:US Postal Service (USPS)	Zip codes
HL7 Vocabulary Domain (Not from PHIN VADS or Searchable File)	Address Type; Address detail data element names eg. county; CPA, street type; STTYPE and Data element names for contacts eg. emergency contact: EC and mobile contact: MC
National Library of Medicine: Medical Subject Headings- NLM MeSH	Age Group eg. frail elderly; infant post mature; child preschool

Table 2. Examples of Demographic Data Elements/Values with No Existing Vocabulary Standard.

Demographic data elements/values
Sex (at birth)(no standard for the data element name, but standards existed for the values)
Transgender
Current Gender (no standard for the data element name, but standards existed for the values)
Occupations(some) eg. prostitute
Birth Type eg. single,twins, >2
Insurance Type(some) eg. CHAMPUS, TRICARE and CHIP
Address Type (some) eg. homeless, correctional facility, foster home
Name Use eg. legal, alias
Response (refused to answer, refused to ask)
Care Taker values: biological parent(s), other relative, foster/adoptive parent, unrelated, and social service

Results

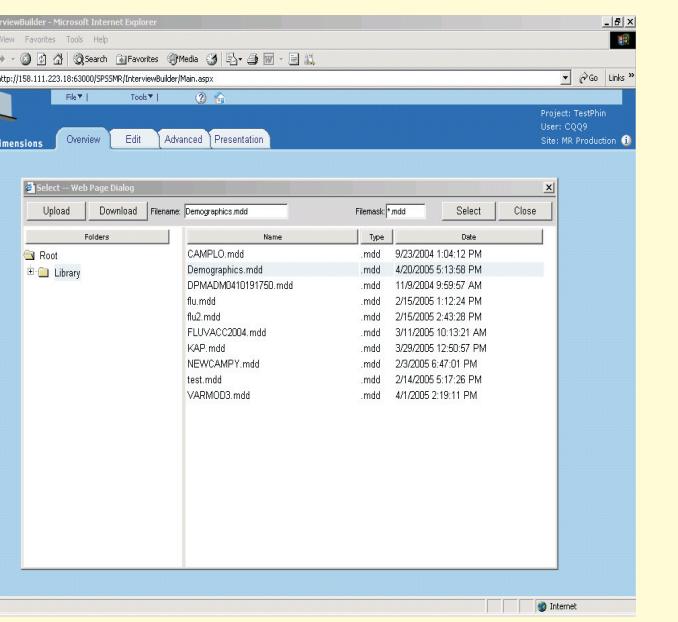


Figure 3. Library of Reusable Demographic Questions (File: Demographics.mdd) Developed and Stored in mrInterview. Note: Some of the Questions had Standard Vocabulary Applied to the Data Elements and Values.

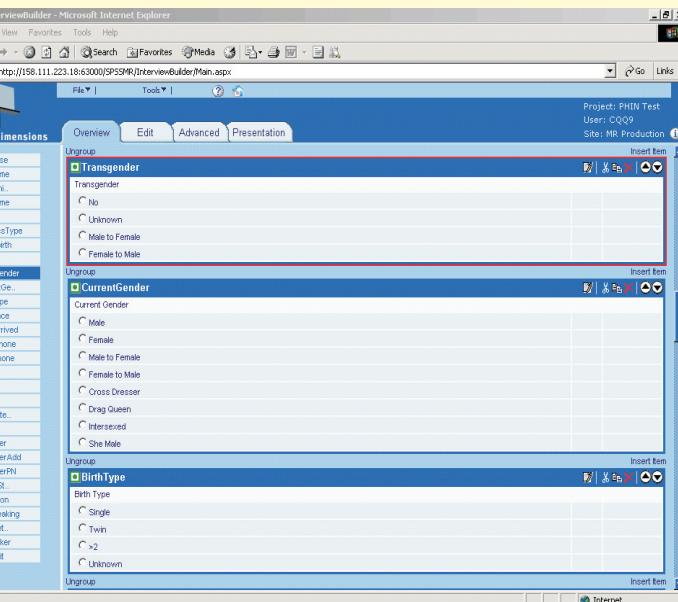


Figure 4. Exhaustive List of Demographic Questions (from the demographic library) Selected From and Applied to a Survey in mrInterview.

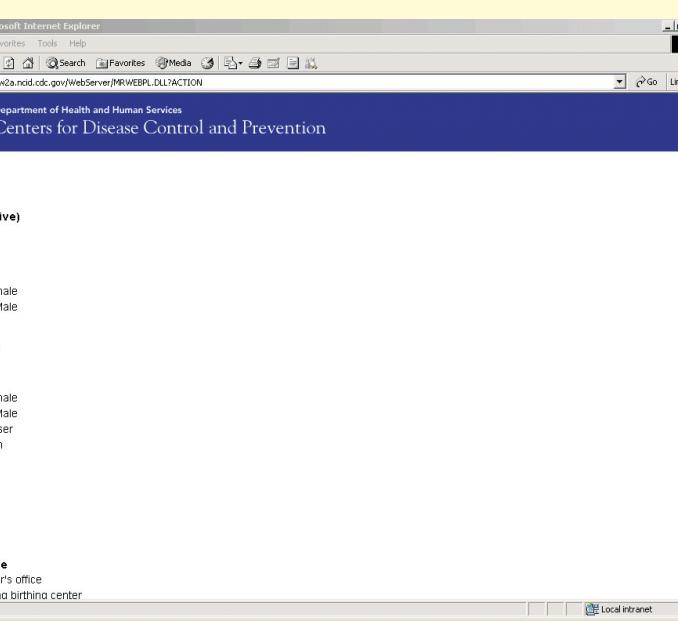


Figure 5. Survey Developed in mrInterview by Applying Reusable Demographic Questions from the Library.

Conclusions

- Most standard vocabularies applied in this project were drawn from PHIN VADS code systems and through the PHIN recommended vocabulary searchable file.
- PHIN VS serves as an appropriate starting point for public health-related vocabulary mapping projects. The PHIN VS tools are most effectively applied when used in consultation with CDC PHIN vocabulary experts.
- Though vocabulary standards exist for majority of the demographic data elements and values, standards that are public health program-specific, are still not available.
- The project exemplified how standard vocabulary could be applied to a tool. The benefit of having applied standard vocabulary to the survey questions, is consistency in interpretation and interoperability between applications once the data is exported.
- The vocabulary standard identification and development effort (including program-specific standards) should engage PHIN vocabulary experts, application developers, as well as program subject matter experts working in close collaboration.

References

1. Public Health Information Network Website, Available at: www.cdc.gov/phin, Accessed (April 5th 2005).
2. Hearings on Uniform Standards for Medical Record Information, Available at: <http://www.ncvhs.hhs.gov/981209t4.htm>, Accessed (April 11th 2005).